

[The Knowledge Bank at The Ohio State University](#)

Article Title: Data for a Database as a General Manuscript Index

Article Author: Hamer, Richard

Issue Date: December 1987

Publisher: William R. Veder, Slavisch Seminarium, Universiteit van Amsterdam,
Postbus 19188, 1000 GD Amsterdam (Holland)

Citation: *Polata Knigopisnaia: an Information Bulletin Devoted to the Study of Early Slavic Books, Texts and Literatures* 17-18 (December 1987): 65-72.

Appears in:

Community: [Hilandar Research Library](#)

Sub-Community: [Polata Knigopisnaia](#)

Collection: [Polata Knigopisnaia: Volume 17-18 \(December 1987\)](#)

DATA FOR A DATABASE AS A GENERAL MANUSCRIPT INDEX

RICHARD HAMER
Christ Church
OXFORD OX1 1DP

INTRODUCTION. There is an obvious place for a comprehensive "union catalogue" database available to all researchers and covering the main categories of manuscript study. Such a database should complement but in no way be a substitute for more specialised database projects, such as those already in existence or in process of creation, just as none of these should become a substitute for the conventional scholarly manuscript catalogue in book form.

This paper contains some thoughts resulting from my own attempt to devise a relational database for use as a general index, whose objects are to facilitate location or listing of manuscripts of various types using, if necessary, multiple criteria (such as "All MSS of works by Bede written in England in the 13th century"), and to direct researchers to other indexes, databases and printed catalogues. (1) Such a database must contain enough basic information to achieve its aims, but should avoid the temptation to become too comprehensive. Other items can, after all, easily be added at a later stage, or the whole or parts of the database can be duplicated for use as the basis of other databases for more specialised purposes.

LIMITATIONS. Though a database can hold a very large amount of material, it is best to avoid excessive size, and one should consider putting manuscripts of widely differing types into separate databases.

I have taken my broad limits as western manuscripts up to about 1500/1550, generally excluding those whose contents were composed after 1500. Later manuscript copies of medieval writings should be included if they have textual significance. I exclude manuscripts written in Greek, Cyrillic, Hebrew, Arabic, and other non-Roman alphabets from my own database, simply because problems of typography and size suggest that it would be better for someone properly qualified to run separate databases, which could be exactly or substantially similar, for such clearly definable categories; there would still of course be some overlap, as with glossed texts and MSS containing both Greek and Latin texts. Equally, in listing a collection of manageable size, it might be best to record all the medieval manuscripts, giving non-Roman or late ones a simple identification and basic references to fuller information elsewhere.

GENERAL CATEGORIES. The categories of information that seem basic are: main codicological facts; contents; scribe and writing; decoration and illustration; earlier history of manuscript; whereabouts of facsimiles, photographs, microfilms; main reference works; special features, such as music.

AUTHORITIES. Where there is possibility of doubt about any statement, such as a date or provenance, warning should always be given, usually

by inclusion of "?". A high proportion of "facts" about manuscripts are in reality "opinions", and the authorities for such opinions must be supplied. It is often obvious who is the most dependable authority, but the existence of disagreements should be recorded in the Comment table.

MAIN CODICOLOGICAL FACTS. Unfortunately it is not always possible simply to use "the manuscript" as the basic unit. Manuscripts become divided, as with the 'Liber custumarum' at the London Guildhall, which consists of parts of two manuscripts bound together; other parts of these two manuscripts are bound together in Oriel College, Oxford, MS 46 and in B.L., MS Cotton Claudius D.ii; entirely separate and unrelated manuscripts can be bound together, or with printed books; they can be written by several people over long periods, as with some Chronicles or Cartularies, and thus defy convenient division into periods; or they can be written at one time and illustrated later, sometimes at another place; they can be written by a continental scribe for an English bishop while moving from country to country, as in the case of some Balliol College manuscripts written for Bishop Gray;(2) and at a simpler level a manuscript may contain corrections or relevant or irrelevant marginal insertions of different dates and provenances. If a researcher is to identify categories by criteria such as date of writing, the records for such manuscripts need to be divided. This must be done in such a way that it is still easy to call up information about the "whole" manuscript. The system that I have adopted at present is to indicate divisions by ending the shelfmark or catalogue-number with "A", "B", etc., (with a space before the capital); this makes it essential that these combinations shall nowhere else be used in the 'Shelfmark' column, so that manuscripts with such shelfmarks as Balliol 173A and 173B must be recorded correctly without the space. Once a MS entry has been divided in this way, no entry relating to it can be made without use of this capital. This raises a slight retrieval problem, as a call for information about "Oriel MS 46" would be greeted with the response that it was not recorded; it would not be hard to devise a formula or program to allow for this, but generally agreed solutions to such problems would be desirable.

SHELFMARK AND NICKNAMES. The Shelfmark/Catalogue Number given is that of a manuscript's last known location, usually its present one; these can be updated when a manuscript moves or is re-discovered (previous locations are recorded in the MS-history Table). A status column reports if a manuscript is Missing or Destroyed or its present whereabouts Unknown. A Nickname column records the popular names ("Lindisfarne Gospels", "Bedford Hours") of certain manuscripts.

COUNTRY. While the exact place of origin of the MS, if known, will be supplied under 'MS History', it is convenient to record country of origin in this section to facilitate general searches. Of course, country of origin is sometimes a matter of opinion, and where there is doubt a "?" should be included. There is also the problem of changing country boundaries and names: what should be done about Flanders, Outremer, Bohemia, for example?

DATES. It is the standard practice, where a precise date is unknown,

to attribute manuscripts to centuries or parts of centuries, using Roman numerals with various appendages indicating first-half, middle, last third, turn of, etc.. Often the cataloguers who use these terms are imprecise about the limits of what they mean; for example, does 'early 14th century' mean that in the opinion of the writer it cannot possibly be after 1310, or 1315, or 1320, or what? Other, or sometimes the same, cataloguers, use terms like "c.1385". Does this mean 'definitely within 5 years either side'? For the computer it is essential to be able to interpret these rough datings with absolute numerical limits to enable precise searches to be made. In view of the rather relative way in which some experts use these terms, I have felt obliged to extend the dates generously in each direction for safety, but this is unsatisfactory. We should put pressure on the experts to state their outside limits precisely.

I have not used the conventional Roman forms of dating, partly because different experts use slightly different versions, but primarily because of the obvious advantage of having Arabic numerals for the computer to handle. I have instead adopted a system whereby any date or range of dates can be allowed for: all dates are given two values, the earliest and the latest possible year, to each of which is attached a reliability column, either blank or containing "?". From this, to find fifteenth-century manuscripts one calls up all whose earliest possible date was before 1501 and whose latest was after 1400. Manuscripts of known date have that year recorded as both earliest and latest.

LANGUAGE, SECFO, RELATED MSS, MAIN AUTHORITY. Other basic pieces of information placed in this Section are: the first words of the Second Folio; the name of the main authority used, which usually means the fullest and most recent catalogue or the best edition; a code for the use of languages other than or in addition to Latin in the MS (a MS in French would be noted as "F", one basically French with some Latin as "F+", and one basically Latin with some French as "+F"); and a Column for closely related MSS, for such cases as MSS which have become divided (as with the parts of Oriel 46 etc. referred to above), for those which are the exemplar of or copied from another surviving MS, and for manuscripts in more than one volume, especially if they have become separated.

The above particulars I regard as fundamental to the central indexing function, and hope that if creators of databases are to co-operate all would agree to supply these to each other, and that as far as possible compatible forms for such entries should be agreed.

TECHNICAL PARTICULARS. A more technical codicological table can record measurements both external and "written space", number of columns and lines, foliation or pagination, material, facts about binding, such as date, name of binder, material. Important as it will be that this information should ultimately be put on a database so that manuscripts of similar format and size can be identified, this will probably have to wait till more fundamental tasks have been completed, and might best be undertaken as a separate project.

TITLES, AUTHORS, TRANSLATORS. The problems of indexing medieval texts

and authors are notorious, given variant titles and incipits, and authors variously known by first or second name, place of origin, Latin and vernacular forms, and so on. These problems could be overcome by the use of an agreed, possibly abbreviated, form or code for every item and author, to be decoded on another table which would relate it to all possible variant forms. Devising these abbreviations and decoding tables would be a major task, and it would be desirable to have it competently done once and for all in a way that will be widely accepted. A possible solution would be to adopt lists of abbreviations already in use in certain authoritative reference works as a basis, modifying them and adding extra ones as needed. Unfortunately such lists as there are often disagree with each other, and in many cases the abbreviations are too long to be satisfactory for computer use. For example, those in works such as the Thesaurus linguae latinae would be suitable for the periods they cover, though some would need to be modified. For some types of work bibliographical references could be used, as with entries in the Index of Middle English Verse or the Bibliographia Hagiographica Latina, or references by volume and page to Migne's Patrologia Latina, though even some of these are inconveniently long. Meanwhile it is easy enough to record major works or broad categories, and the final ideal form, in which all sermons, versions of chronicles, individual saints' lives, charters, etc., are identifiably recorded, must lie a long way in the future. The practice of noting the contents of MSS by broad type, such as medical, legal, Statutes, astronomical, etc., is useful in itself, as well as a convenient preliminary expedient, and will be worth retaining as a separate column.

Translators should be recorded as well as authors, and allotted their own separate column.

SCRIBES AND SCRIPTS. Where a scribe is anonymous and has no accepted designation (like "the tremulous hand of Worcester") one will have to be devised, perhaps based on a MS or a locality. It would be desirable to record in general terms the type of writing, as Humanistic, Carolingian Minuscule, Bastard Anglicana, etc., but there will be problems with terms which are not yet generally agreed.

ILLUMINATIONS. Much the same applies to the Illuminators Table as to the Scribes; foliation of all decorations and illustrations should be supplied, and different categories noted, as Historiated or other forms of decorated Initial, Sketches, Illustrations, Borders, etc..

MS HISTORY. MS History will record all known earlier locations of the MS with dates, including exact place of origin if known, and sales. Signatures and names which do not necessarily indicate ownership would also be noted here.

FILMS AND FACSIMILES. One of the most important functions of this database will be to record the existence and whereabouts of any sort of photographic reproduction. By providing this information it is hoped to help preserve manuscripts themselves from unnecessary exploratory investigation.

a repository for noting problems and for recording important categories of information not dealt with elsewhere, such as the presence of musical notation, mottoes, heraldic devices and the like.

RELATIONS WITH OTHER DATABASE MAKERS. Databases designed for more specialised purposes will necessarily include many of the above items, and it is hoped that their creators would be willing to make them available. It is not suggested that anyone should supply or offer information on categories outside his own particular task; and above all no-one should attempt to supply opinions on matters outside his own expertise (an anxiety often expressed by scholars with reservations about the use of computers for this sort of purpose).

Since there is so much to be done in putting MS information onto databases, any move towards co-operation and the avoidance of duplication of labour is to be welcomed. In particular there would be great advantage in achieving as much agreement as possible on the forms of entry to be used for many types of information; while this could be largely done by negotiation in many cases, such as forms for shelfmarks, symbols for languages, etc., there is obviously a need for a concerted programme of action in matters like the naming of texts and authors. Agreement on representation of accents, cedillas and umlauts is another desideratum, and, no doubt more controversially, for the use of a common language for certain types of entry.

Footnotes:

(1) A short account of my database, "MSS", appears in Gazette du livre médiéval, 9 (Automne 1986), pp.7-11. Among the numerous databases and catalogues I have studied, I am particularly indebted to the various publications of the Institut de Recherche et d'Histoire des Textes, and to the published account of the Benjamin Data Bank by N.L.Hahn and J.B.Smith. L.Burnard of the Oxford University Computing Service has technically devised "MSS" and is responsible for many refinements and simplifications to my own proposals.

(2) R.A.B.Mynors, Catalogue of the manuscripts of Balliol College Oxford, (Oxford, 1963), pp.xxvi et seq.

TABLES FOR DATABASE "MSS"

MS A single identifiable ms

*NUMBER	I	Unique serial number of ms in database
MSCOL	8	Short name for ms collection, = COLLECTION.CTAG
CATNO	24	Catalogue number or brief name
STATUS	1	Code for current status (lost, destroyed, etc)
ORIGIN	8	Country of origin
DATE-A	I	Year of earliest possible date
DATE-A-R	1	Reliability of DATE-A
DATE-B	I	Year of latest possible date
DATE-B-R	1	Reliability of DATE-B
DATER	8	Authority for dating
SECFOL	32	First words on second folio
REL	16	Close relationship to other ms (by MS.NUMBER)
NICKNAME	24	Other name attached to ms (e.g. Bedford Hours)
OTHERLANG	4	Code for presence of language(s) other than Latin

MSDESC Technical details of a ms

*MSNO	I	= MS.NUMBER
FORP	1	Foliated, paginated or neither (F, P or N)
FOLIOS	32	Foliation (or number of folios)
HEIGHT	I	Height in mm
WIDTH	I	Width in mm
WSHIGH	I	Written space height in mm
WSWIDE	I	Written space width in mm
NCOLS	1	Number of columns
NLINES	6	Number of lines in page
MATERIAL	8	Material of ms if not vellum
BINDING	12	Binder/rough date of binding

COLLECTION Library or other collection of mss

*CTAG	8	Unique code for ms collection (= MS.MSCOL)
NAME	32	Full name of collection
TOWN	16	Location of collection
COUNTRY		Country

TITLE Any work or item appearing in one or more mss

*TTAG	12	Short title to identify the item
AUTH	36	Author
INCIPIT	80	Incipit or name
TYPE	1	Verse, Prose, Mixture
SUBJ	4	Code for type of subject matter
LANG	6	Code for language(s) in item

MS-CONTENTS Particulars of version of an item in one ms

*MSNO	8	MS.NUMBER
*TTAG	12	as in TITLE table
*FOLS	12	Foliation in the particular ms
VERSION	12	Any comment on this version

SCRIBE Details of an identifiable scribe

 *SCRTAG 8 Unique tag to identify a scribe
 SCRNAM 32 Name or other description of individual scribe

 MS-SCR-DIV A A scribal subdivision in a ms

 *MSNO I = MS.NUMBER
 *SCRDIV 2 Code for scribal division in ms (A,B etc)
 FOLIOS 12 Foliation of this division
 SCRID 8 Tag for scribe (= SCRIBE.SCRTAG)
 SCRIPT 24 Description of script
 SCRFUNC 2 Scribal function, main, gloss, rubric, corrector, etc

 PUBLICATION Reference or other work

 *SHTIT 12 Short title of work
 TITLE 80 Full title
 AUTHOR 80 Author or editor
 PUBDATE 8 Date of publication

 MS-REF Details of published reference to a ms

 *MSNO I MS.NUMBER
 *PUBN 12 PUBLICATION.SHTIT
 DETAIL 12 Page number or other identification within publication

 MS-FACS Whereabouts of reproductions

 *MSNO I MS.NUMBER
 *FOLIOS 12 Folios reproduced
 *PUBN 12 = PUBLICATION.SHTIT
 OWNER 8 Identity of owner or collection
 FACTYPE 8 Type of facsimile, microfilm, etc

 DIALECT Dialect particulars

 *DICODE 4 Tag to identify the dialect
 DINAME 76 Description of dialect (e.g.area)

 MS-DIAL-DIV Dialectal division of a ms

 *MSNO I = MS.NUMBER
 *DIADIV 2 Dialectal division as scribes (A, B etc)
 FOLIOS 12 Folios containing the division
 DIATAG 4 = DIALECT.DICODE

 ILLUMINATOR Particulars of an illuminator

 *TAG 8 Code to identify the illuminator
 NAME 32 Name/nickname of illuminator
 ATELIER 16 Name of atelier, centre, etc

MS-ILL-DIV Details of illuminations of a ms

*MSNO	I	= MS.NUMBER
*FOLIOS	32	Folios on which illuminations appear
ILLUM	32	Types of illumination
ARTIST	8	= ILLUMINATOR.TAG

MS-HIST Previous owners and locations

*MSNO	I	MS.NUMBER
*OWNER	8	Owner or institution or sale
CATNO	24	Catalogue number or sale date
DATE	12	Specific dates or century
PRMK	24	Previous pressmark or other catalogue number

COMMENT Assorted matter not in other tables

*MSNO	I	= MS.NUMBER
CTYPE	2	Code for type of comment
CTEXT	80	Comment on Heraldry, Music, Watermark etc

TIT-REF Published reference to or edition of title

*TTIT	12	= TITLE.TTAG
*SHTIT	12	PUBLICATION.SHTIT

An asterisk before a column name means that it is an identifier, and an entry must be made. The second column gives the maximum length of entries; "I" in this column means that entries must consist of whole numbers.

Other tables can be added, and none of the details above is unalterable; in fact alterations in the light of experience are generally quite easy to introduce.